

# The Cryo-EM structure of AAV2 Rep68 in complex with ssDNA reveals a malleable AAA<sup>+</sup> machine that can switch between oligomeric states

Vishaka Santosh<sup>1</sup>, Faik N. Musayev<sup>2</sup>, Rahul Jaiswal<sup>1</sup>, Francisco Zárata-Pérez<sup>1,3</sup>,  
Bram Vandewinkel<sup>4</sup>, Caroline Dierckx<sup>4</sup>, Molly Endicott<sup>1</sup>, Kamyar Sharifi<sup>1</sup>, Kelly Dryden<sup>5</sup>,  
Els Henckaerts<sup>4</sup> and Carlos R. Escalante<sup>1,\*</sup>

<sup>1</sup>Department of Physiology and Biophysics, Virginia Commonwealth University School of Medicine, Richmond, VA, USA, <sup>2</sup>Department of Medicinal Chemistry, School of Pharmacy, Virginia Commonwealth University School of Medicine, Richmond, VA, USA, <sup>3</sup>The Massey Cancer Center, Virginia Commonwealth University School of Medicine, Richmond, VA, USA, <sup>4</sup>Department of Cellular and Molecular Medicine and Department of Microbiology, Immunology and Transplantation, KU Leuven, Leuven, Belgium and <sup>5</sup>Department of Molecular Physiology and Biological Physics, University of Virginia School of Medicine, Charlottesville, VA, USA

Received September 01, 2020; Revised November 01, 2020; Editorial Decision November 03, 2020; Accepted November 27, 2020

## ABSTRACT

The adeno-associated virus (AAV) non-structural Rep proteins catalyze all the DNA transactions required for virus viability including, DNA replication, transcription regulation, genome packaging, and during the latent phase, site-specific integration. Rep proteins contain two multifunctional domains: an Origin Binding Domain (OBD) and a SF3 helicase domain (HD). Studies have shown that Rep proteins have a dynamic oligomeric behavior where the nature of the DNA substrate molecule modulates its oligomeric state. In the presence of ssDNA, Rep68 forms a large double-octameric ring complex. To understand the mechanisms underlying AAV Rep function, we investigated the cryo-EM and X-ray structures of Rep68–ssDNA complexes. Surprisingly, Rep68 generates hybrid ring structures where the OBD forms octameric rings while the HD forms heptamers. Moreover, the binding to ATP $\gamma$ S promotes a large conformational change in the entire AAA<sup>+</sup> domain that leads the HD to form both heptamer and hexamers. The HD oligomerization is driven by an interdomain linker region that acts as a latch to ‘catch’ the neighboring HD subunit and is flexible enough to permit the formation of different stoichiometric ring structures. Overall, our studies

show the structural basis of AAV Rep’s structural flexibility required to fulfill its multifunctional role during the AAV life cycle.

## INTRODUCTION

The adeno-associated virus (AAV) non-structural Rep proteins drive a variety of DNA transactions required for the AAV life cycle including, DNA replication, transcription regulation, site-specific integration, and genome packaging (1–7). Remarkably, all these processes are carried out primarily by two functional domains shared by four Rep proteins. The two large Rep proteins (LReps) Rep78 and Rep68, participate in all AAV DNA transactions while the only attributed biological role for the two small Rep proteins (sReps), Rep52 and Rep40, is during genome packaging where they play a central role as the molecular motors (8). The LReps share two domains: an N-terminal origin binding domain (OBD) and an SF3 helicase domain (HD) joined by a linker of ~25 residues (Figure 1A); additionally, Rep78 has a Zn-finger domain at the C-terminus that is not involved in DNA binding (4,6,9–11). The sReps consist of the helicase domain, but Rep52 also has the Zn-finger domain (4). As members of the SF3 family of helicases, all AAV Rep proteins have a version of the AAA<sup>+</sup> ATPase domain that lacks the classical C-terminal sensor 2 domain but instead have an N-terminal helical bundle known as oligomerization domain (OD) (Supplementary Figure S1) (12,13). Studies have shown that the OBD has

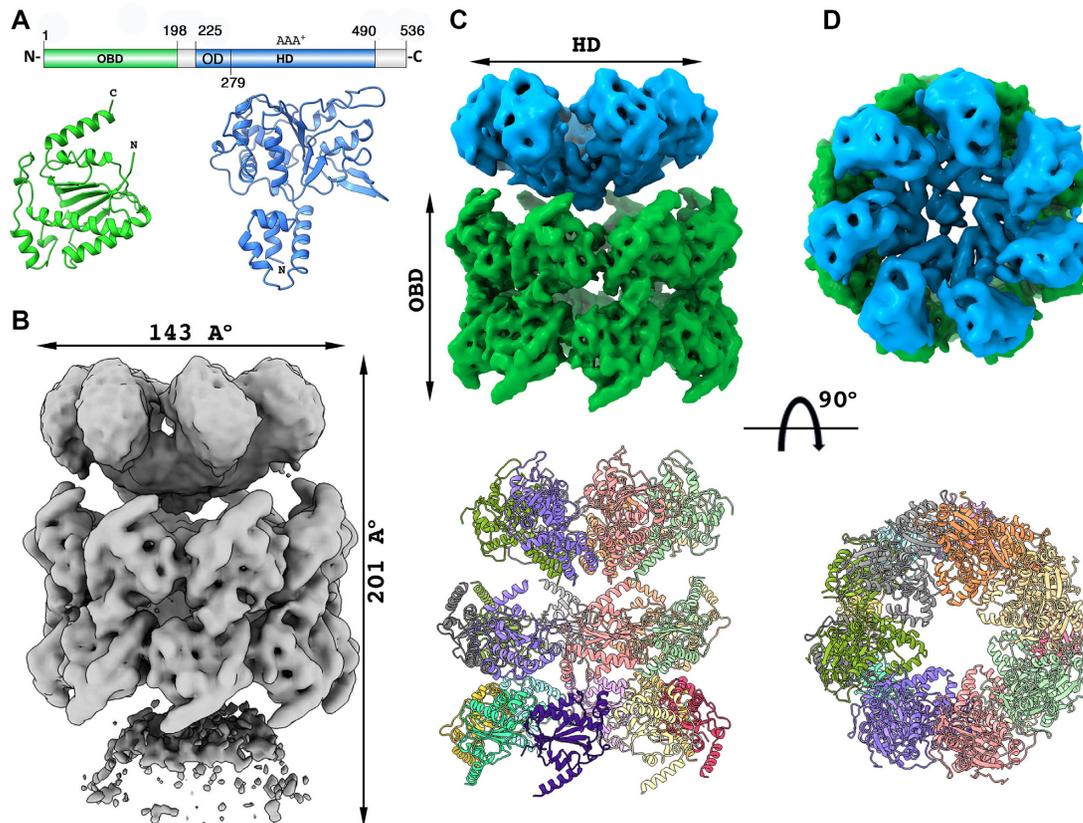
\*To whom correspondence should be addressed. Email: carlos.escalante@vcuhealth.org

Present addresses:

Vishaka Santosh, CCDC CBC Biodefense Branch, MD, USA.

Molly Endicott, University of the West of England Bristol, Department of Applied Sciences, Bristol, UK.

Kamyar Sharifi, Undiagnosed Diseases Network, National Human Genome Institute Clinical center, Frederick, MD, USA.



**Figure 1.** Overview of the Rep68-ssDNA DOC. (A) Primary and X-ray structures of Rep68 functional domains. Color scheme is maintained throughout all figures. OBD: origin binding domain, green; HD: helicase domain, blue; OD: oligomerization domain. (B) 3D Cryo-EM map of the Rep68-ssDNA DOC and overall dimensions. (C, D) Top panel: Composite local cryo-EM maps of HD1 (blue) and OBD DOC (green), Bottom panel: Ribbon representation of the corresponding atomic models. X-ray structures of AAV-2 OBD (PDB:5BYG) and AAV-2 Rep40 (PDB: 1S9H) were used to generate all atomic models.

three non-overlapping DNA binding motifs that carry out distinct functions: A HUH-fold catalyzes the cleavage of ssDNA in the terminal resolution reaction during DNA replication (14–16); a dsDNA binding pocket recognizes Rep binding sites (RBS) found in the origin of replication, p5 promoter site and AAVS1 integration site (11,15,17); and a third motif that binds ssDNA hairpins (14,15,18). The helicase domain has an additional DNA binding motif known as the pre-sensor 1  $\beta$ -hairpin (PS1 $\beta$ H), common in all SF3 helicases and involved in melting and unwinding (Supplementary Figure S1) (12,19). Taken together, the LReps have multiple DNA binding elements that cooperate to catalyze all AAV DNA transactions.

Although the overall domain structure of the AAV Rep proteins is similar to those of other viral SF3 helicases, they could be considered a subclass of this helicase family because of their distinctive oligomerization properties. Whereas common viral SF3 helicases and most AAA<sup>+</sup> proteins form only hexameric rings, Rep40-containing only the helicase domain-is a monomer in solution and only forms transient dimers in the presence of ATP (16,20,21). In contrast, the LReps display a dynamic association process forming a spectrum of oligomeric species, from dimers to octameric rings (20). Furthermore, the interdomain linker plays a critical role in the oligomerization of the LReps as its substitution by a non-related linker, completely abol-

ishes the ability of Rep68 to oligomerize (22). Therefore, the minimum AAV Rep helicase domain -as represented by Rep40- is oligomerization deficient while oligomerization in the LReps requires the cooperative interaction of multiple domains producing complexes that are intrinsically dynamic. More importantly, this conformational plasticity allows LReps to acquire different quaternary structures modulated by the nature of the DNA site. For instance, binding to the integration site AAVS1 produces a heptameric complex, whereas binding to poly-dT ssDNA results in the formation of a double-octamer (23). To understand the structural basis of these processes, we determined cryo-EM structures of Rep68 in complex with ssDNA. The structures reveal that each functional domain can form oligomers with different stoichiometries. Thus, while the OBD favors the formation of octamers, the HDs forms heptamers. Upon binding ATP $\gamma$ S, the HDs transition into hexameric rings. Moreover, the structures show that the key structural motif inducing HD oligomerization is a linker region of 11 residues that acts as a flexible latch docking into neighboring OD domains and stabilizing both heptamers and hexamers. Our results provide essential information that explains the remarkable conformational plasticity of AAV Rep proteins that makes them a unique subclass of SF3 helicases, illustrate the versatility of the AAA<sup>+</sup> domain, and offer insights into their DNA remodeling functions.

## MATERIALS AND METHODS

### Cloning and mutagenesis of rep expression constructs

All mutant proteins were generated using the pHisRep68/15b plasmid which contains the AAV2 Rep68 ORF (1–536) or the truncated form (1–490) subcloned in vector PET-15b (Novagen). Site-directed mutagenesis for all mutants was done using the QuickChange<sup>®</sup> mutagenesis kit (Stratagene). The sequences of all constructs were confirmed by DNA sequencing (GeneWiz).

### Protein expression and purification

All proteins were expressed using the pET-15b vector, in *Escherichia coli* BL21(DE3) cells (Novagen), and purified as described before (23). Rep68\*, a C151S mutant was used in all experiments and will be referred as Rep68 in the manuscript (20). The final buffer contains (25 mM Tris–HCl [pH 8.0], 200 mM NaCl, and 2 mM TCEP). His6-PreScission Protease (PP) was expressed in BL21(DE3)-pLysS at 37°C for 3 h, in LB medium containing 1 mM IPTG. Cell pellets were lysed in Ni-Buffer A (20 mM Tris–HCl [pH 7.9 at 4°C], 500 mM NaCl, 5 mM Imidazole, 10% glycerol and 1 mM TCEP). After five 10-s cycles of sonication, the fusion protein was purified using a Ni-column—equilibrated in Ni-buffer A. Protein eluted was desalted using buffer A and a HiPrep<sup>™</sup> 26/10 desalting column (GE Healthcare). The hexahistidine tag was removed by PreScission protease treatment using 150 µg PP/mg His-PP-Rep68 at 4°C. Subsequent Ni-column chromatography using buffer B (same as buffer A but with 1 M imidazole), was performed to remove the uncleaved fusion protein, and untagged Rep68 was eluted with 50 mM imidazole. Rep68 was finally purified by gel filtration chromatography using a HiLoad Superdex 200 16/600 PG column (GE Healthcare) and Size Exclusion buffer (25 mM Tris, 200 mM NaCl, 1 mM TCEP, pH 8.0). N-terminus His6-tagged WT and mutant Rep68 proteins were concentrated to 2 mg/ml with 50% glycerol, flash-frozen in liquid N<sub>2</sub>, and stored at –80°C.

### Gel-filtration experiments

Gel-filtration studies were performed with the Rep68 concentration held constant at 30 µM and the DNA held constant at 10 µM. The gel-filtration buffer used was: 10 mM Na<sub>2</sub>HPO<sub>4</sub> pH 7.0, 150 mM NaCl, 1 mM TCEP and all results were plotted using Origin (Origin labs). The experiments shown in Supplementary Figure S2 were conducted on a Superose 6 10/300 GL column (GE Healthcare) while the experiment shown in Figure 2 was conducted on a Superose 6 Increase 10/300 gel filtration column (GE Healthcare).

### qPCR-based replication assay

293T cells were transfected with pAV2 or mutant pAV2 using polyethylenimine (PEI, Polysciences). Two hours post-transfection, cells were infected with adenovirus at a multiplicity of infection (M.O.I.) of 5. The medium was replaced with fresh DMEM + 10% FCS 6 h post-infection. The cells were harvested when a clear cytopathic effect was

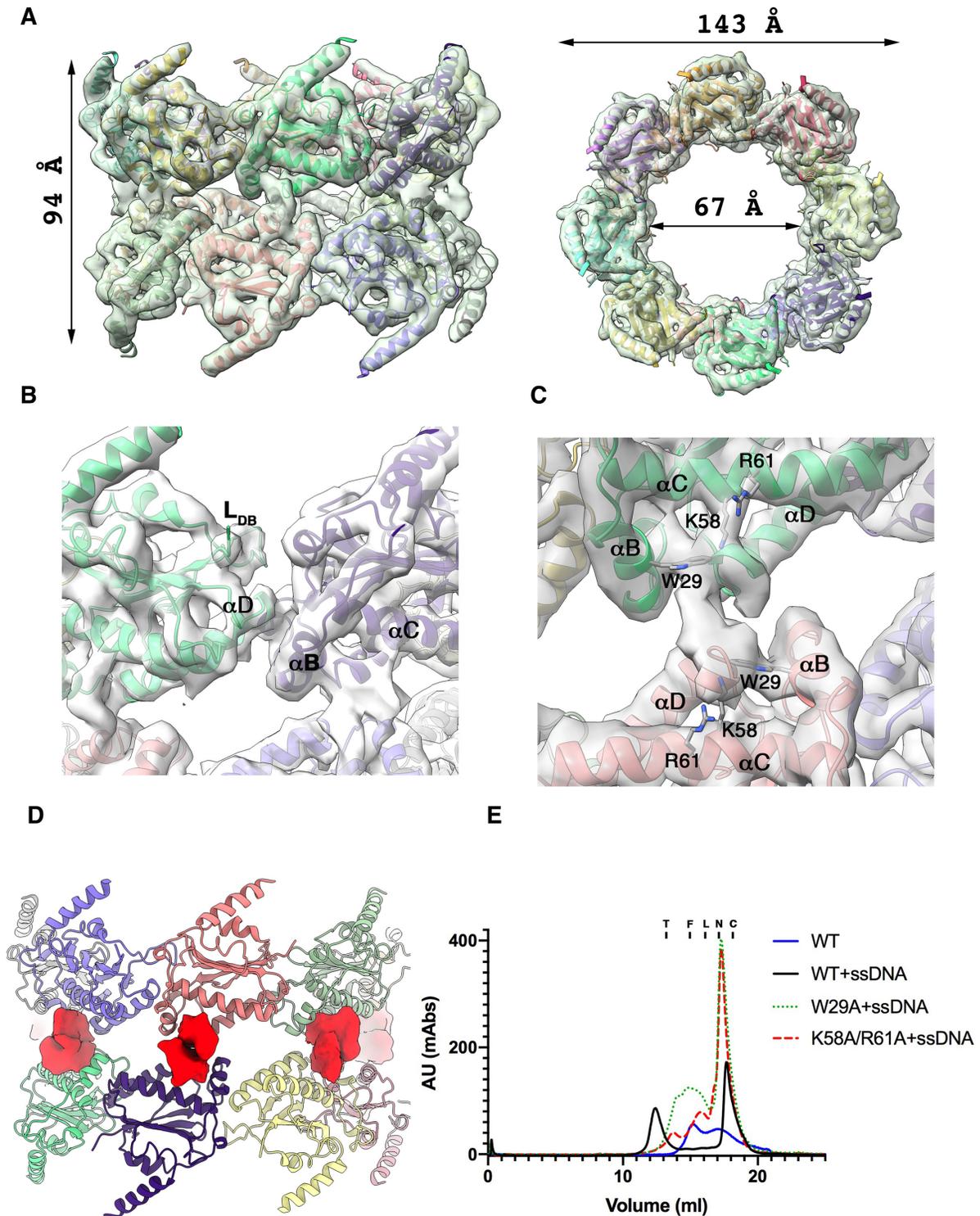
visible (~45 h post-infection). Total DNA was extracted from the cells using The GenElute<sup>™</sup> Mammalian Genomic DNA Miniprep Kit from Sigma Aldrich. The samples were diluted to 4 ng/µl and 5 µl was used for quantification with forward (fw) and reverse (rv) primers which bind to the cyclophilin gene (fw primer: TGCTGGACCCAACACAAA TG; rv primer: TGCCATCCAACCACTCAGTCT) and primers which bind to *cap* (fw primer: TTCTCAGATGCT GCGTACCGGAAA; rv primer, TCTGCCATTGAGGT GGTACTTGGT). As master mix, the PowerUp<sup>™</sup> SYBR<sup>™</sup> Green Master Mix from Applied Biosystems was used. The cycling conditions were the following: a 2-min initial heating step at 50°C and a 5-min initial denaturation step at 95°C, followed by 39 cycles of a denaturation step at 95°C for 30 s, an annealing step at 58°C for cyclophilin primers and 60°C for cap primers for 20 s, and an elongation step at 72°C for 1 min.

### AAV Infectious particles assay

Two parallel systems were used. In one, AAVpro<sup>®</sup> 293T (Takara) cells were transfected in duplicates using the calcium phosphate method with an AAV2 ITR-containing plasmid including a GFP gene (pTRUF11), a plasmid expressing AAV2 Rep WT or mutants (pRC2) and Cap, and third plasmid containing the adenovirus helper proteins (Takara). 48 hours post transfection, cells were harvested and lysed using AAV extraction solution A & B (Takara) as per the manufacturer's recommendation. HT1080 cells were infected with increasing amounts of viral solution and the percentage of GFP-positive cells was determined after three days of infection. In the second protocol, HEK293T cells were triple transfected in a 6-well format using polyethylenimine (PEI) Max with 2 µg of an AAV2 ITR-containing plasmid including GFP as transgene, 1.6 µg of a helper plasmid expressing AAV2 Rep (WT or mutant) and Cap, and 1.6 µg of a third construct containing the adenovirus helper functions (pXX6, University of North Carolina Vector Core Facility). Seventy-two hours post-transfection, 1 ml of the supernatant was spun down to clear the cellular debris and increasing volumes of supernatant were used to transduce HeLa cells at 60–70% confluency. Before transduction, the medium of the HeLa cells was replaced with DMEM + 5% FCS. 48h after transduction, the percentage of GFP-positive HeLa cells was determined by flow cytometry (Accuri<sup>™</sup> C6 Plus; BD Biosciences) (24).

### Analytical ultracentrifugation

Sedimentation velocity experiments were carried out using a Beckman Optima XL-I analytical ultracentrifuge (Beckman Coulter Inc.) equipped with a four and eight-position AN-60Ti rotor. Protein samples (420 µl; final concentration 0.25 mg/ml) were prepared in a buffer containing 25 mM Tris–HCl, pH 8.0, 200 mM NaCl and 1 mM TCEP. Samples were loaded and centrifuged in 2-sector carbon-filled Epon centerpieces at 25 000 rpm. In all experiments, temperature was kept at 20°C. Sedimentation profiles were recorded using UV absorption (280 nm) and interference scanning optics. For the analysis of the results the program Sedfit was used to calculate sedimentation coefficient distribution profiles using the Lamm equation (25).



**Figure 2.** Structure of OBD double-octamer. (A) Two orthogonal views of the local Cryo-EM maps and 16 docked OBDs. (B) Ribbon representation of the OBD-OBD interface in the octamer. The  $L_{DB}$  (DNA binding loop) and  $\alpha D$  interact with  $\alpha B$  and  $\alpha C$  residues in the neighboring subunit. Both motifs also recognize the GTC repeats in RBS DNA sites. (C) Ribbon representation of the interface generating the DOC. Residues W29, K58 and R61 have been identified in previous studies as critical for ssDNA binding. The positions of the side chains are from the docked X-ray model. (D) Unaccounted density (red) located between OBD molecules in the crevice found between  $\alpha B$ -C and  $\alpha D$ . (E) Size exclusion chromatography profiles of W29A and K58A/R61A mutants and their ability to form the DOC. Samples were run on a Superose 6 increase column, with Rep68 WT-ssDNA DOC eluting at  $\sim 12$  ml (black line), W29 mutant (green dotted) elutes at 15 ml and K58A/R61A at 16 ml. Rep68 is shown in blue and ssDNA elutes at 18 ml. Protein markers shown correspond to thyroglobulin (T, 669 kDa); ferritin (F, 440 kDa); aldolase (L, 158 kDa); conalbumin (N, 75 kDa) and carbonic anhydrase (C, 29 kDa).

### Helicase assay

The substrate used in this assay is a heteroduplex DNA consisting of an 18-bp duplex region with a 15-nucleotide 3' tail at the bottom strand. The top strand has the sequence 5'-AGAGTACGGTAGGATATGAACCAGACACATGAT-3'; the bottom strand with sequence 5'-CATA TCCTACCGTACTCT-F-3' is labelled at the 3' end with fluorescein and is released upon unwinding. We used the unlabeled bottom strand as a trap to prevent reannealing of the displaced fluorescent strand. All reactions were performed in a buffer containing 25 mM HEPES, 50 mM NaCl (pH 7.0) in a total volume of 50  $\mu$ l. Rep68 at different concentrations was mixed with double stranded F-DNA at a final concentration of 1  $\mu$ M and incubated for 15 min. Reaction was started by adding 5 mM ATP-Mg and 2.5  $\mu$ M trap DNA. Reaction was incubated at 25°C for 1 min. EDTA was used to stop the reaction at a final concentration of 20  $\mu$ M. Aliquots of 10  $\mu$ l were loaded in a 12% bis-acrylamide gel (30%) (19:1) using 6X-loading dye (0.25 xylene cyanol FF, 30% glycerol). A Gel Doc EZ Imager was used to conduct densitometry and analysis. Background lane subtraction, white illumination and an activation time of 300 sec was used for the analysis.

### EM sample preparation

The Rep68-dT<sub>25</sub> complex was prepared as previously described (23). In short, Rep68 was mixed with a 25-mer poly-dT oligonucleotide (dT<sub>25</sub>), purified to homogeneity and concentrated in the presence of 0.05% n-octyl- $\beta$ -D-glucopyranoside (OG). The Rep68-AAVS1-15 ssDNA complex was made by incubating 30  $\mu$ M protein with 15  $\mu$ M ss AAVS1-15 DNA in reaction buffer (10 mM Na(PO<sub>4</sub>)<sub>2</sub> pH 7.0, 150 mM NaCl 1 mM TCEP) for 15 min at room temperature. The complex was then concentrated to 200  $\mu$ l using Amicon Ultra-4 centrifugal filter (Millipore) and loaded onto a pre-equilibrated Superose 6 Increase 10/300 gel filtration column (GE Healthcare). Eluted complex was concentrated back to 1/8th and 1/16th of original protein concentration. C-Flat carbon grids CF1.2/1.3-4C, 400 mesh Cu (Electron Microscopy Sciences CF413-50) were glow-discharged for 45 seconds with amyamine using a PELCO easiGlow™ glow-discharge system. Just before spotting the sample, 5 mM MgCl<sub>2</sub> and 5 mM ATP <sub>$\gamma$</sub> S (final concentration) was added to the complex and incubated at room temperature for 5 min. 3.5  $\mu$ l sample was spotted, blotted and plunged frozen into liquid ethane using the Vitrobot unit and stored in liquid nitrogen.

### Cryo-EM grid preparation and data collection

C-flat grids were glow-discharged for 40 seconds with amyamine using a PELCO easiGlow™ glow-discharge unit and spotted with 3.0  $\mu$ l of a 0.5 mg/ml Rep68-dT<sub>25</sub> sample. The sample was manually blotted for 1.5 s and plunged into liquid ethane. For storage, the samples were stored in liquid nitrogen. To obtain more side views, 0.05% octyl-beta-glucoside (OG) was added to the sample and frozen as previously described (23). Initial screening was done on a Tecnai F20 electron microscope equipped with a 4k  $\times$  4k ultrascan CCD camera and preliminary data collection on a

Titan Krios at the University of Virginia Molecular Electron Microscopy core facility. Final data sets for the DOC were collected at the National Center for CryoEM Access and training (NCCAT) located at the New York Structural Biology Center on a Titan Krios in counting equipped with a Gatan K2 Summit direct detector. For the apo-data, 40 movie frames were recorded at 165 000 magnification and total dose of 92.02 electrons per  $\text{\AA}^2$  at a pixel size of 0.85  $\text{\AA}$ . For the ATP <sub>$\gamma$</sub> S complex, the samples were collected at 22 500 $\times$  at a pixel size of 1.073  $\text{\AA}$  and total dose of 55.46 electrons/ $\text{\AA}^2$ . Data for the Rep68-ssAAVS1-ATP <sub>$\gamma$</sub> S were collected at the Pacific Northwest Center for Cryo-EM (PNCC) on a Titan Krios equipped with a Falcon 3 direct detector. Samples were collected at 59 000 $\times$  magnification at a pixel size of 1.415  $\text{\AA}$  and total dose of 50 electrons/ $\text{\AA}^2$ .

### Cryo-EM image processing and model building

*Rep68-dT<sub>25</sub> (DOC)*. Micrographs frames were aligned and superimposed using Motioncor2 (26). Contrast transfer function (CTF) parameters were calculated with CTFFind4. 2D classification of  $\sim$ 1000 manually selected particles were used for template-based particle picking in RELION v2.0 (27,28). Selected particles (118 531) were imported into cryoSPARC v.2.3 and multiple rounds of 2D classification were done to remove false particles and junk classes. A final set of 62 512 were extracted and used to generate four *ab initio* 3D models. The particles for the best 3D model particles were again used to generate two *ab initio* models that were further processed using heterogeneous refinement. The final model was obtained using non-uniform refinement to an overall resolution of 4.6  $\text{\AA}$  using the 0.143 FSC cutoff. C1 symmetry was applied during the entire refinement process. To improve the density of the HD1 portion of the map, local refinement was performed with a mask encompassing the entire HD1 heptameric ring with signal subtraction for both OBD and HD2. The resulting map had a nominal resolution of 6.7  $\text{\AA}$  had better overall density, particularly in the OD region where the main protein chain could now be identified. Local refinement was also performed on the OBD double-octameric ring with particle subtraction of the HD1 and HD2 regions. The resulting map had an overall resolution of 4.4  $\text{\AA}$ .

*Rep68-dT<sub>25</sub>/ATP <sub>$\gamma$</sub> S*. Preprocessing was done in Relion starting with 4555 micrographs using Motioncor2 and CTFFIND4. After importing into cryosparc, a set of 294,324 particles were visually inspected through a series of passes of 2D classification and 2D selection to remove junk and low-quality particles, resulting in 78,701 good particles. After generating an initial *ab initio* model, and 3D non-uniform refinement the final map had an overall resolution of 5.2  $\text{\AA}$  using the FSC 0.143 cutoff. We performed local refinement in the HD1 portion of the map with a mask encompassing the entire HD1 heptameric ring with signal subtraction for both OBD and HD2. The resulting map had a nominal resolution of 7.4  $\text{\AA}$ .

*Rep68-ssAAVS1/ATP <sub>$\gamma$</sub> S*. All steps were performed in cryoSPARC starting with 2669 movies. The movie frames were aligned and summed using Motioncor2 to obtain the

final dose-weighted images. Estimation of the CTF was performed using CTFFIND4. In total 907,616 particles were selected from a total of 2584 images. 2D classification resulted in a mixture of hexameric and heptameric classes, that based on inspection of the top views correspond to about 63.1% and 36.9% of the total particles respectively. Reconstruction of the heptameric complex started with 248,270 particles that were subjected to multiple rounds of 2D classification and then two *ab initio* models were generated. The models were further analyzed to exclude false-positive and contaminant particles through iterative cycles of heterogeneous classification and 2D classification to generate a final stack of 150,263 particles. A final reconstruction was determined using non-uniform refinement to an overall resolution of 5.2 Å. We used 3D variability analysis to characterize conformational heterogeneity of the complex.

**Hexameric Rep68-dT<sub>25</sub> complex.** Grids were prepared with Rep68-dT<sub>25</sub> complex as described previously. Data analysis started with 1722 movies corrected with MotionCor2 and CTFFIND4 for CTF estimation. Automatic particle picking was performed in RELION to generate an initial stack with 578,679 particles. After 2D classification in cryoSPARC a mixture of particles representing double-octameric rings (45.5%), heptameric (8.9%) and hexameric (45.6%) could be identified. For the reconstruction of the hexameric complex, 185,495 particles were passed through multiple cycles of 2D classification and iterative heterogeneous classification to generate high quality classes for final reconstruction. Final cryo-EM map was calculated using 61,195 particles to an overall resolution of 5.01 Å. We performed local refinement with signal subtraction of the vestigial OBD density. The map had better density in the region of the OD ring although the overall resolution was 5.1 Å. 3D variability analysis was used to characterize the conformational heterogeneity of the complex.

All molecular models were built with Chimera, Coot and refined with Phenix (29–31). The crystal structure of Rep68 AAV-2 OBD (PDB:5BYG) was used to build the double-octameric model. Initially, 16 individual OBDs were manually docked into the cryo-EM map and oriented using Fit in Map in Chimera. The resulting model of the double-octameric OBD was further refined using rigid body refinement in PHENIX. A similar approach was used to build the apo heptameric HD using the crystal structure of AAV-2 Rep40 (PDB:1S9H). For the ATP $\gamma$ S complexes, during initial refinement in PHENIX, two rigid bodies were defined for each helicase subunit. The rigid body 1 contained the OD (residues 1–277) and rigid body 2 the AAA<sup>+</sup> domain (residues 278–490). All superposition analysis of the models was done using the program LSQMAN (32). Electrostatic surface potential calculation using the docked X-ray Rep40 structure was done using Pymol (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC).

### Crystallization and X-ray structure determination

The oligonucleotides used for crystallization were purchased from Integrated DNA Technologies, Inc. Site A: 5'-ATATGCTCGCTCTT-3'. DNA was purified on a

MonoQ-5/50GL column. The purified DNA was desalted, lyophilized, and resuspended in TE buffer (10 mM Tris-HCl, pH 8.0, 40 mM NaCl, 1 mM EDTA). OBD and DNA were mixed in a 1:1.1 molar ratio. Complex was concentrated to a final concentration of protein of about 20 mg/ml. The buffer concentration was exchanged during the concentration process to 25 mM Tris-HCl, pH 7.5, 100 mM NaCl and 1 mM TCEP. Crystallization experiment was carried out using both hanging- and sitting-drop methods with commercially available screening kits at 4°C. Crystals grew from 3  $\mu$ l hanging drop after 2–3 weeks. The best crystals were obtained from reservoir solution containing 10 mM Na cacodylate, pH 6.5, 30% PEG400 and 0.2 M LiCl. The crystals belonged to the space-group I222 with unit cell dimensions  $a = 170.2$  Å,  $b = 173.2$  Å,  $c = 173.5$  Å. Diffraction data was collected at the National Synchrotron Light Source (NSLS) at the Brookhaven National Laboratory beamline X6a. The data was processed with the program iMosfilm (CCP4 suite) (33). The structure was solved by molecular replacement using the program PHENIX (31). We used the structure of the AAV2 OBD as a search model (pdbid 4ZQ9). Model building was carried out using PHENIX and manual building using the program COOT (30).

## RESULTS

### Overall architecture of the Rep68-ssDNA complex

To gain insights into the molecular mechanisms underlying the multifunctionality of AAV Rep proteins and their unique oligomeric properties, we carried out a combination of cryo-EM and X-ray studies of several Rep68-ssDNA complexes. First, we focused on our previously characterized Rep68-ssDNA double-octameric complex, from here on referred to as DOC (23). Using single-particle cryo-electron microscopy, we were able to obtain a cryo-EM map at an average resolution of 4.6 Å (Figure 1, Supplementary Figure S2, Table 1). The structure of the complex resembles a double funnel with overall dimensions of 201 Å  $\times$  133 Å  $\times$  143 Å ( $H \times W \times D$ ). The map can be divided into three ring-like sections: a well-defined central ring composed of OBDs (RING<sub>OBD</sub>) flanked by two outer helicase rings (HD1, HD2). The density of HD2 is too weak and fragmented to identify individual subunits (Figure 1B). Likewise, the densities of the linkers connecting the OBDs to the HDs are not visible, indicating that they are highly dynamic. A close examination of the RING<sub>OBD</sub> map shows that it consists of two octameric OBD rings arranged in a head-to-head orientation (Figure 1C). Surprisingly, there is a symmetry mismatch between the OBD and HD rings, with HD1 showing seven well-defined domains (Figure 1C, D). Altogether, the results show that the DOC has a central stable core of two octameric OBD rings and two highly dynamic HDs rings.

### The double-octamer is stabilized by ssDNA

The OBD double ring is 94 Å wide with an outer diameter of 143 Å generating an internal channel with a diameter of  $\sim$ 67 Å (Figure 2A). The cryo-EM map of the

**Table 1.** Cryo-EM data collection, refinement and validation statistics

	Apo-(DOC)		ATP $\gamma$ S-(DOC)	Heptamer-ssAAVS1-ATP $\gamma$ S	Hexamer-dT <sub>25</sub> -ATP $\gamma$ S
Detector	Gatan K2 Summit		Gatan K2 Summit	Falcon 3	Gatan K2 Summit
Magnification	165 000 $\times$		22 500 $\times$	59 000 $\times$	22 500 $\times$
Voltage (kV)	300		300	300	300
Electron exposure (e <sup>-</sup> /Å <sup>2</sup> )	92.03		55.46	50.0	50.0
Defocus range ( $\mu$ m)	-1.0 to -2.2		-1.5 to -3.0	-2.4 to -1.1	-1.25 to -2.75
Pixel size (Å)	0.825		1.073	1.415	1.073
Symmetry imposed	C1		C1	C1	C1
# images	1449		4555	2669	1722
Initial particles	118 531		294 324	433 139	578 679
Final particles	46 031		78 701	193 869	61 195
Map resolution (Å)	4.6		5.2	5.2	5.0
FSC threshold	0.143		0.143	0.143	0.143
<b>Refinement*</b>					
Cryo-EM Map and model	OBD-DOC 7JSE, EMD-22451	HD1 7JSF, EMD-22452	HD-Heptamer 7JsG, EMD-22453	HD-Heptamer 7JsH, EMD-22454	HD-Hexamer 7JSI, EMD-22455
Initial models used (PDB code)	5BYG	1S9H	1S9H	1S9H	1S9H
Map resolution (Å)	4.4 (local)	6.7 (local)	7.4 (local)	5.2	5.1 (local)
FSC threshold	0.143	0.143	0.143	0.143	0.143
<b>Model composition</b>					
Non-hydrogen atoms	13288	7815	7963	7923	6630
Protein residues	3168	1932	1904	1904	1656
DNA	32	79	339	300	160
<i>B</i> factors (Å <sup>2</sup> )					
Protein	156.5	69	43	43	68.5
DNA	22.6	20	20	20	143
<b>r.m.s.d.</b>					
Bond lengths (Å)	0.001	0.004	0.001	0.004	0.003
Bond angles (°)	0.381	1.134	0.277	1.104	0.424
<b>Validation</b>					
Molprobtity score	1.25	1.18	1.99	1.18	1.30
Clash Score	0.41	0.2	5.66	0.00	0.59

\* All protein models contain main chain atoms only unless specified.

RING<sub>OBD</sub> is of enough resolution to fit the complete main-chain backbone from the OBD X-ray structure and some of the bulkier side chains (residues 1–198) (Figure 2A). Sixteen individual OBDs generate a double-octameric OBD ring model. The only region from the X-ray structure that did not fit entirely into the cryo-EM density after the initial docking was the C-terminal  $\alpha$ -helix F, which is slightly tilted by  $\sim 2^\circ$ . The OBD molecules are oriented such that the HUH catalytic residues reside in the inside of the ring. Surprisingly, the interface involved in generating the octameric rings utilizes the same motif that recognizes the dsDNA RBS GCTC repeats, namely, helices D and the loop L<sub>DB</sub> (Figure 2B) (17). This region interacts with  $\alpha$ -helices B, C and the loop connecting strands  $\beta 2$  and  $\beta 3$  in the neighboring OBD molecule (Figure 2B). In all, this interface potentially buries 1180.5 Å<sup>2</sup> of solvent accessible area for a total of  $\sim 4720$  Å<sup>2</sup>/octamer (17).

A second interface accounts for the interactions mediating the formation of the double-octameric ring and comprise a pocket generated by  $\alpha$ -helices B, C and D in each of the sixteen OBD molecules (Figure 2C). Interestingly, this region contains three residues (W29, K58, R61) that were previously identified as critical to binding ssDNA hairpin RBE' found in the AAV5 ITR (15). A closer inspection of the Cryo-EM map around this pocket, shows an unaccounted density that spans neighboring OBDs across in-

dividual octameric rings (Figure 2D). We hypothesize that the unaccounted density corresponds to ssDNA bridging OBD rings. To test this hypothesis the predicted interacting residues (W29, K58 and R61) were mutated to alanine and were tested for their ability to form the complex. Figure 2E shows that any of the mutants failed to form the complex. To understand the DNA structural requirements that lead to the formation of the Rep68 DOC, we carried out size-exclusion chromatography experiments with different DNA substrates. First, we aimed to identify the minimal length of ssDNA required to form the complex. The elution profiles of each sample show that ssDNA smaller than 17 nucleotides failed to form the complex (Supplementary Figure S3A). Next, we explored the effect of the composition of the ssDNA sequence using poly-guanine, poly-cytosine, and a region from the integration site AAVS1. Elution profiles show that the complex forms more readily using ssDNA molecules that are pyrimidine-rich, with preference dT > dC > dG (Supplementary Figure S3B).

### The X-ray structure of the OBD with ssDNA shows formation of double-octameric rings

In previous work, we showed that Rep68 forms octameric rings, which represent one among an assortment of oligomers present in solution (20). We hypothesized that

the OBD by itself may assemble as an octameric ring under the right conditions. To test this hypothesis, we performed sedimentation velocity experiments at different concentrations and compared the experimental sedimentation coefficient to the sedimentation coefficient calculated for different OBD atomic models, including an OBD octameric ring using the program SOMO (34,35). Results show that at concentrations below 124  $\mu\text{M}$ , the OBD forms two different species, one sedimenting at  $\sim 2.2$  S and a faster species at  $\sim 5$  S (Supplementary Figure S4A). Higher concentrations result in additional species with a sedimentation coefficient of 7.2 S, which is similar to that predicted for the octameric ring (7.3 S) (Supplementary Figure S4A). In addition, we also performed glutaraldehyde crosslinking of OBD, showing the formation of octamers on an SDS-PAGE (Supplementary Figure S4B). These experiments demonstrate that the formation of octameric rings is an intrinsic property of the OBD.

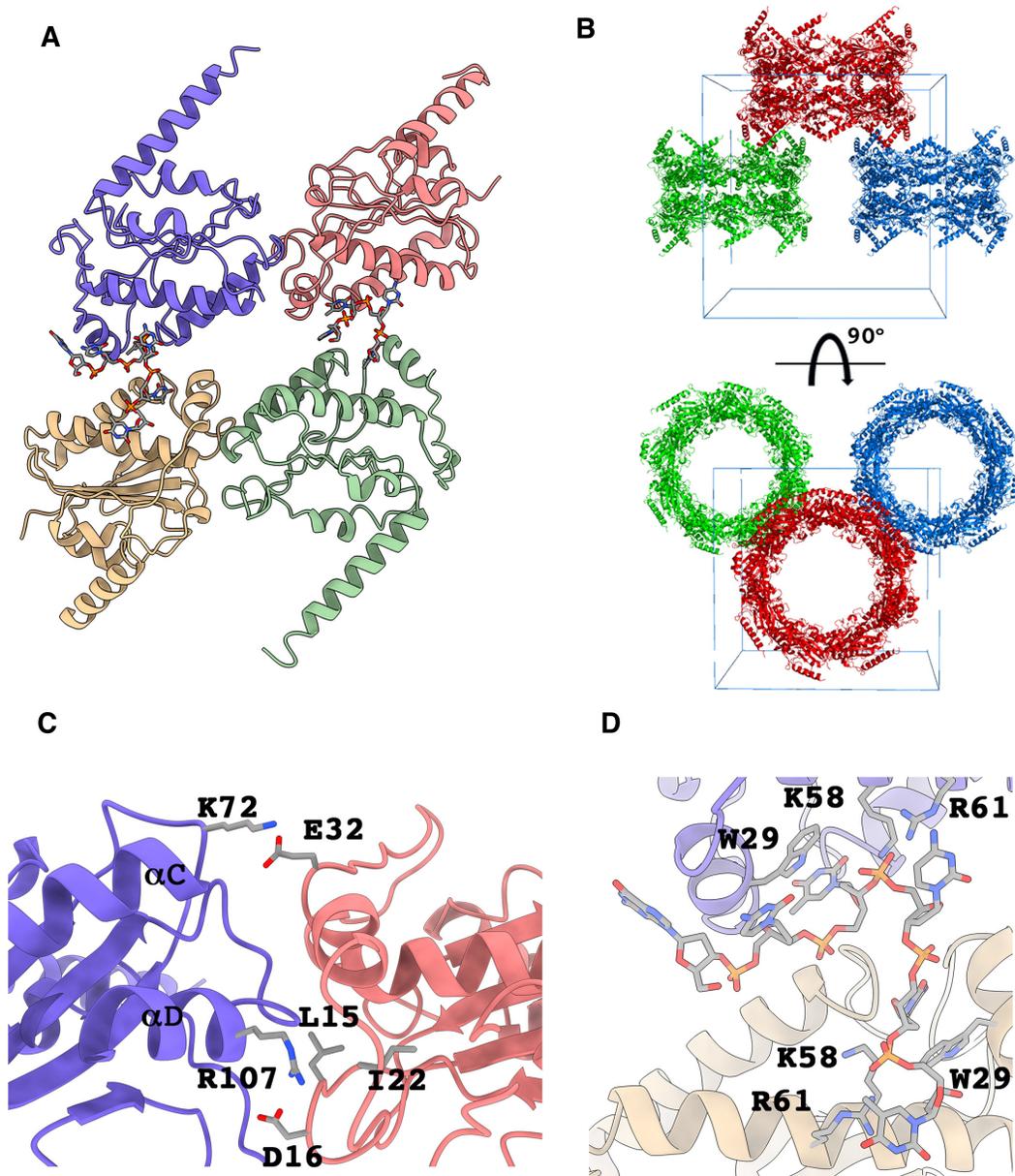
We determined the crystal structure of OBD in complex with a ssDNA molecule containing an RBS sequence from the AAV2 ITR. The complex crystallizes in space group I222 with four OBD molecules in the asymmetric unit sharing two molecules of ssDNA (Figure 3A, Table 2). The arrangement of the four OBD molecules resembles that of the cryo-EM Rep68-DOC using the same OBD-OBD interface motifs with ssDNA bridging two molecules (Figure 3A). Moreover, analysis of the arrangement of the molecules in the crystal lattice shows that the OBD assembles as double octameric rings (Figure 3B). The dimensions of this ring are identical to the cryo-EM structure such that a model of the octameric ring from the crystal structure can be manually docked into the cryo-EM density producing an almost perfect fit. Examining the protein contacts in the X-ray structure, we identified several interface residues forming the octameric ring. An example of such interactions occurs between  $\alpha$ -helix C residue K72, forming a salt bridge with E32 on helix B of the neighboring OBD molecule (Figure 3C). Of particular interest is residue R107 in helix D that forms multiple interactions with L15, D16 and I22 of the neighboring OBD molecule (Figure 3C). Not surprisingly, an R107A mutation is defective in Rep68-ssDNA complex formation, as we had previously shown in the initial Rep68-DOC characterization study (23). The crystal structure also shows how the DOC assembles by the ssDNA bridging of two octameric rings. The DNA used in the crystallization was a 14-mer with sequence 5'-ATATGCTCGCTCTT-3' that adopts a stem-loop conformation. A pair of OBD molecules bind to the 3'-half of the ssDNA molecule interacting with the last five nucleotides (CTCTT). The sequence TCT interacts across two OBD molecules with each thymine making stacking interactions with residue W29 (Figure 3D) and forming a cation- $\pi$  interaction with R122 in each OBD. Other key residues from this region are K58 and R61 that make phosphate backbone interactions with the middle cytosine of this tri-nucleotide region. This interface is equivalent to the one found with AAV5 OBD in complex with the RBS' stem-loop and suggests that the OBD ssDNA binding motif has a strong preference for pyrimidine rich DNA (15). Moreover, it explains, as previously noted, how mutations of residues W29, R58, and K61 abolish the formation of the Rep68-DOC.

### Heptamerization of the helicase domain is directed by the OD

The Rep68 HD1 region shows seven well-defined densities forming a two-layered ring. The smaller tier ring is made up of the N-terminal oligomerization domains (ODs), forming a continuous density with an external diameter of 75  $\text{\AA}$  (Figure 4A). The larger tier with a diameter of 134  $\text{\AA}$  is discontinuous, with well-separated densities consisting of seven AAA<sup>+</sup> domains (Figure 4A). The resolution of the ODs is similar to that of the OBDs and we can trace the complete main chain (residues 224–277). In contrast, the resolution of the AAA<sup>+</sup> domain is significantly lower. 3D variability analysis shows that the HD1 ring is very dynamic, moving as a rigid body with respect to the core OBD octamer, and also each individual helicase domain moving independently, particularly the AAA<sup>+</sup> subdomain (Supplementary Figure S5). The overall HD1 ring has a funnel-like inner channel with a decreasing radius from 76  $\text{\AA}$  at the top of the AAA<sup>+</sup> ring to 25  $\text{\AA}$  at the lower end of the OD ring (Figure 4A, right panel). Surprisingly, the HD1 model shows that heptamerization is only mediated by interactions between the ODs. The X-ray Rep40 structure can be fitted into each HD density without any major changes in its overall conformation. The cryo-EM density also indicates that the N-terminal  $\alpha$ -helix 1 in the OD can be extended with 4–6 residues from the linker region (Figure 4B). Based on the X-ray docked model, the OD-OD interface potentially buries 536  $\text{\AA}^2$  of surface area and is made up of helix 1 and 4 in one subunit and helices 2 and 3 in the neighboring OD (24). Although the cryo-EM map is not of sufficient resolution to reveal side chain interactions, the fitted X-ray atomic model shows multiple potential interface interactions between OD subunits, particularly N254 is in close proximity to N269 and Y224 of a neighboring subunit, while S261 and R260 may also interact (Figure 4C). To test this potential interface, we mutated both residues to alanine and measured the effect on viral replication in an infectious particle assay. Figure 4D shows that each mutation causes a dramatic decrease in the production of infectious AAV particles implying the predicted interface residues are functionally significant.

### The OD contains a second DNA interacting motif

There is an additional density at the center of the HD1 heptameric ring structure that likely represents part of a ssDNA molecule (Figures 1A and 4A). The size of the density fits a ssDNA of  $\sim$ four nucleotides (Figure 4E). Surprisingly, the DNA density is not located around the pre-sensor 1  $\beta$ -hairpin (PS1 $\beta$ H) but resides instead at the N-terminal end of the OD ring. Based on the X-ray docked model, several basic residues such as R260 and K264 generate an electropositive surface area (Figure 4F, G). The presence of ssDNA at this location is a reflection of the larger electropositive charge in this region as compared to the PS1 $\beta$ H ring (Figure 4G). Due to the low resolution of the cryo-EM map, we do not know the actual positions of the side chains around this area, but we can make predictions from the X-ray structure and test if these residues play a role in AAV life cycle. We mutated R260 and K264 to alanine and performed an infectious particle assay (24). Figure 4H shows that both mutations completely abolish the formation of infectious AAV particles indicating they play a critical func-



**Figure 3.** X-ray Structure of OBD-ssDNA complex. (A) Overall view of the asymmetric unit with four OBD molecules bridged by two ssDNAs. (B) Two orthogonal views of the crystal lattice showing double octameric assemblies. (C) Ribbon representation of the OBD-OBD octameric interface and some of the most representative residues involved in protein-protein interactions. (D) Structure of the bridging ssDNA showing stacking interactions of W29 with thymines and K58 and R61 interacting with the phosphate backbones.

tional role during the AAV life cycle and failed to replicate AAV DNA (Supplementary Figure S6). However, helicase assays show that although R260A and R264A mutants can unwind DNA, they have lower activity when compared to the WT implying they are not essential to unwind DNA (Figure 4I).

#### **In the presence of ATP $\gamma$ S, the AAA<sup>+</sup> domain undergoes a major conformation change**

To learn about the structural changes occurring upon ATP binding, we solved the cryo-EM structure of Rep68 with two ssDNA substrates in presence of ATP $\gamma$ S. Interest-

ingly, during particle picking and 2D classification, we observed the presence of hexameric ring particles in both cases (Supplementary Figure S7). The 2D classes of the DOC-ATP $\gamma$ S complex show the HD rings with lower density than the OBD region (Supplementary Figure S8). The DOC-ATP $\gamma$ S map was refined to an overall resolution of 5.2 Å (Figure 5A, Supplementary Figure S8) and we performed local refinement of the HD1 map region to a final resolution of 7.4 Å. The cryo-EM map of the Rep68-ssAAVS1 was solved to 5.2 Å and in contrast to the DOC complex, we observed no density for the OBDs, suggesting they are highly mobile (Figure 5B, Supplementary Figure S9). The HD cryo-EM maps in both complexes are less defined than

**Table 2.** X-ray data collection and refinement statistics

	OBD-ssRBS
Space group	I222
Cell dimensions	
<i>a</i> , <i>b</i> , <i>c</i> (Å)	170.2, 173.2, 173.5
$\alpha$ , $\beta$ , $\gamma$ (°)	90, 90, 90
Wavelength (Å)	0.9795
Resolution (Å)	40.86–3.30 (3.418–3.30)
No. of measured	363 016
No. of unique	38 846
Data coverage (%) <sup>a</sup>	99.9 (100)
$R_{\text{merge}}$ (%) <sup>a,b</sup>	0.125 (0.594)
$1/\sigma^a$	15.5 (4.3)
<b>Refinement statistics</b>	
Resolution range	40.86–3.30
Reflections	38830
$R_{\text{work}}$ (%) <sup>c</sup>	20.12
$R_{\text{free}}$ (%) <sup>d</sup>	21.81
Non-hydrogen atoms	6689
Protein	6549
DNA	136
Metal	4
Water	0
Average <i>B</i> -factors (Å <sup>2</sup> )	72.12
Protein	70.84
DNA	83.59
Metal	145.43
Water	0
Rms deviations	
Bonds (Å)	0.005
Angles (Å)	0.82
Ramachandran plot quality	
Favored (%)	94.99
Allowed (%)	4.63
Outliers (%)	0.38

<sup>a</sup>Values for the outmost shells are given in parentheses.

<sup>b</sup> $R_{\text{merge}} = \sum |I - \bar{I}| / \sum I$ , where *I* is the integrated intensity of a given reflection.

<sup>c</sup> $R_{\text{work}} = \sum ||F_o| - |F_c|| / \sum |F_o|$ .

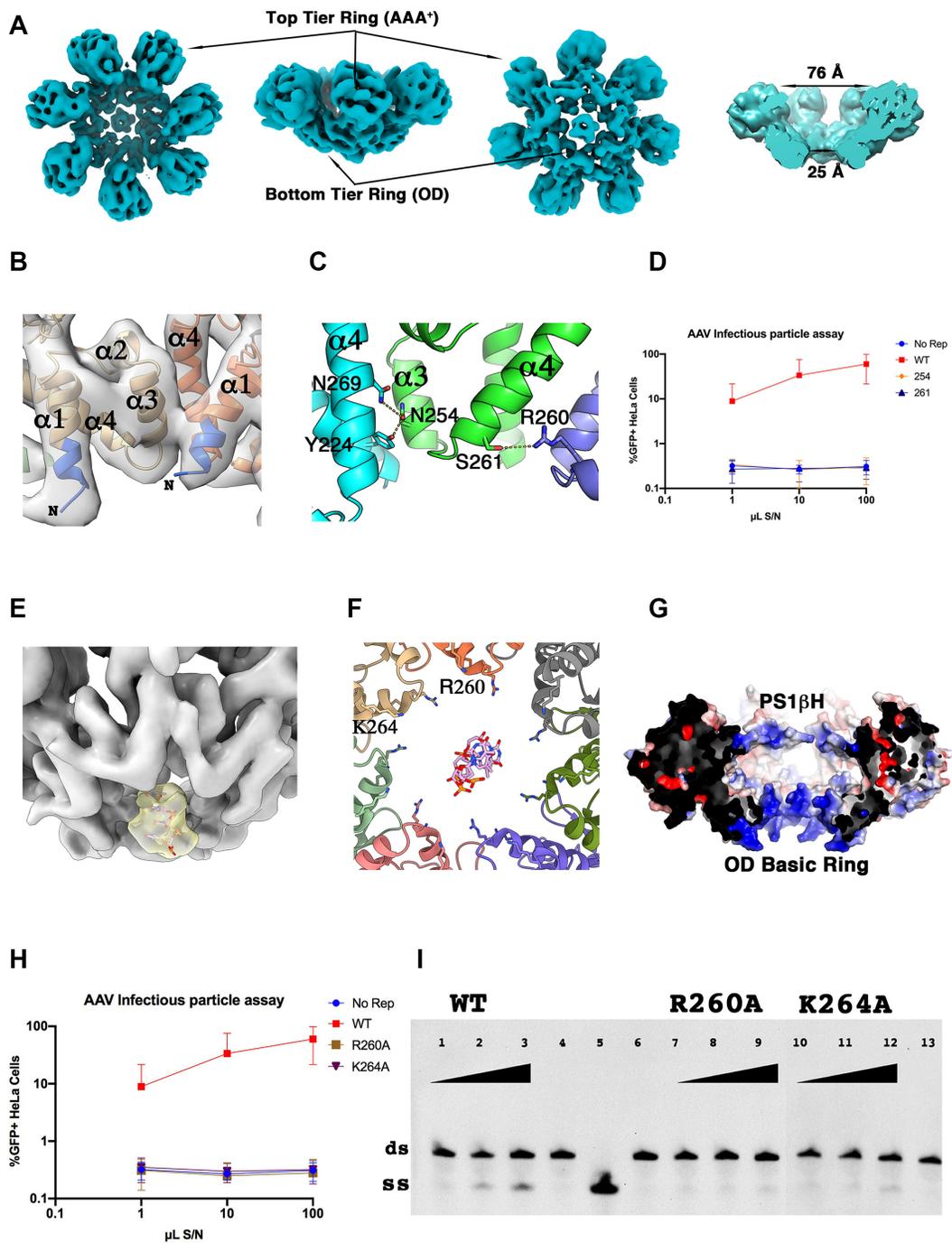
<sup>d</sup>For *R* free calculations 5% of data was excluded from refinement.

in the apo structure, with the ODs forming a continuous density that is incomplete in some regions, suggesting a very dynamic complex. A combination of factors may account for the lack of definition of the HD cryo-EM densities, including the presence of multiple conformations. We undertook a 3D variability analysis using cryoSPARC v2.15 to determine the conformational heterogeneity of one of the heptameric complexes (36,37). Results show that the heptamer exists in several distinct major conformations that include open, closed and rings divided into halves of four and three subunits, illustrating the dynamic character of the structure (Figure 5C and Supplementary Movie S1). For the final reconstruction, classes that clearly showed open or divided rings were not incorporated in the final 3D refinement. Although low resolution, the maps are sufficient to obtain an overall fit of the helicase domains and determine gross conformational changes. Initially, docking of the X-ray Rep40 structure did not produce a good fit, instead, the ODs and AAA<sup>+</sup> domains were fitted as separate rigid bodies. The resulting model shows a large rigid body rotation of the AAA<sup>+</sup> domains with respect to the ODs of ~44° toward the center of the ring, causing the diameter at the top of the inner channel to constrict from 76 to 46 Å (Figure 5D). The ssDNA density spans the length of the cen-

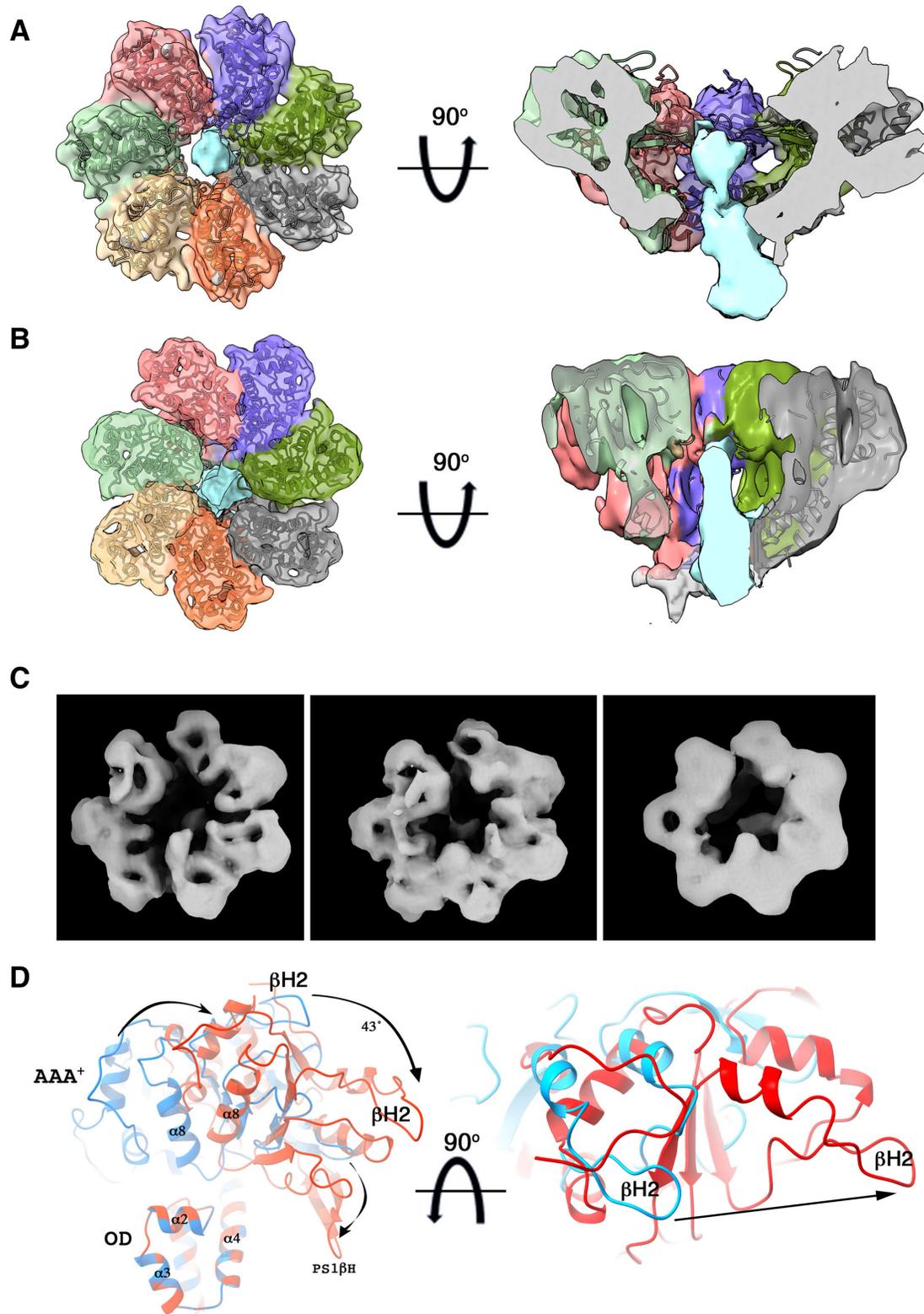
tral channel and protrudes at the lower OD-tier. In both structures, the ssDNA shows bending around the OD region. Two to three subunits are clearly contacting DNA through residues of the OD ring and the AAA<sup>+</sup> domain, although the resolution is not sufficient to see specific side chains. Taken together, the ATP $\gamma$ S-bound structures show that upon binding ATP $\gamma$ S, Rep68 undergoes a large conformational change that induces AAA<sup>+</sup> ring closure and potential DNA deformation.

### The hexameric complex shows that the linker forms a flexible latch that stabilizes the oligomeric rings

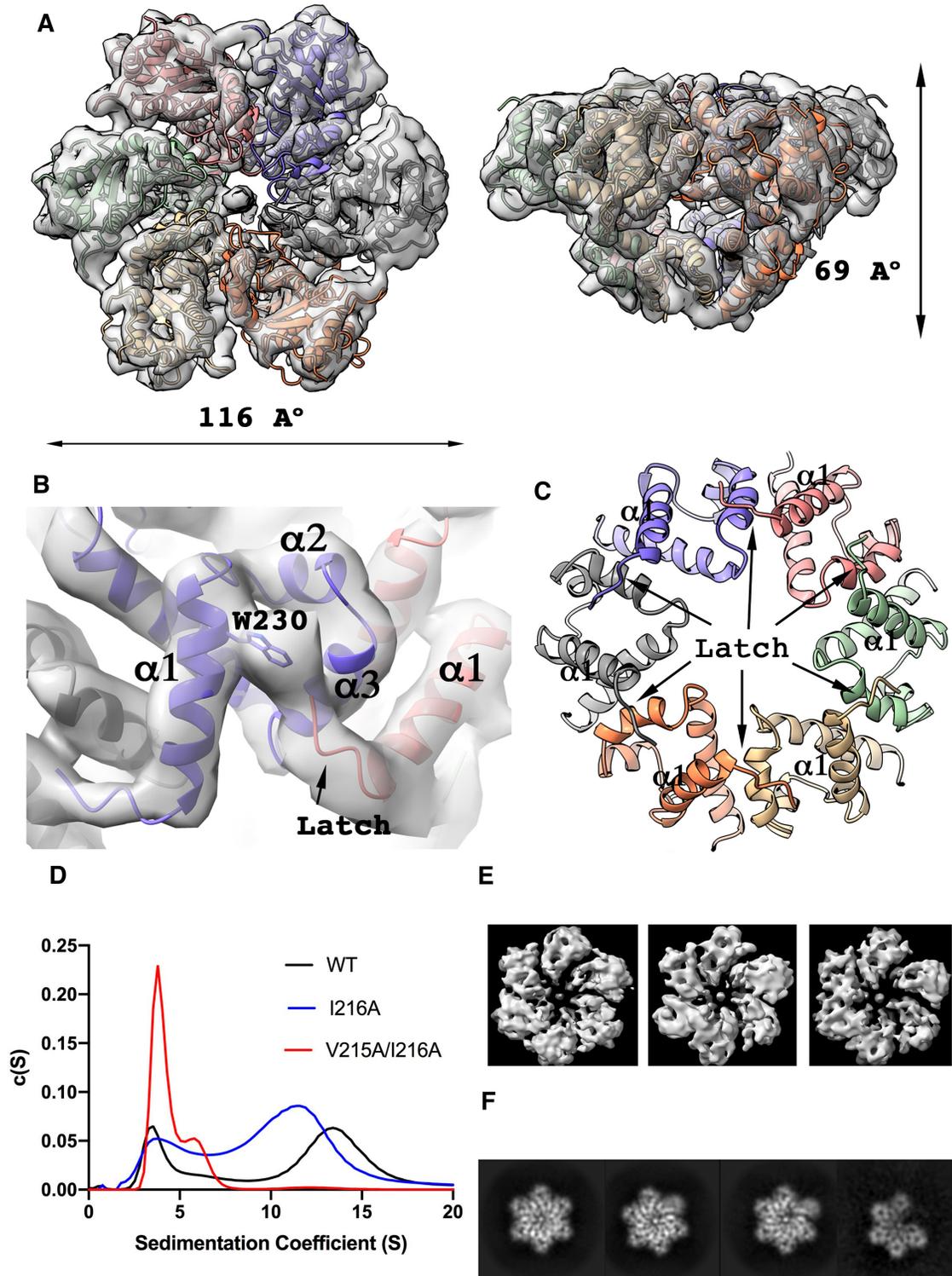
We next determined the cryo-EM structure of the Rep68-dT<sub>25</sub> hexameric ring complex as the hexameric-ssAAVS1 complex did not have enough particle orientations to generate a reasonable model. The structure was solved to an overall resolution of 5.1 Å after local refinement (Figure 6A, Supplementary Figure S10). The overall arrangement of the hexamer resembles that of other SF3 family members, but with significant differences. Particularly, the arrangement of the subunits does not follow a proper 6-fold rotational symmetry, displaying differences in how each subunit docks into the ring. The diameter of the ring is 116 Å in width and 69 Å in height with an inner tunnel that has narrowed to ~13 Å. The rigid body rotation of the AAA<sup>+</sup> domain with respect to the apo structure is ~40°, similar to the heptameric ATP $\gamma$ S complex. Most of the observed density for ssDNA resides along the region of the PS1 $\beta$ H region with some minor density around the entrance of the OD. The cryo-EM map shows additional unaccounted density that further extends the N-terminal end of the OD to include five additional residues from the linker region. Part of this region docks into a crevice formed between OD helices 1 and 3 of the neighboring molecule in a clockwise direction (Figure 6B, C). We hypothesize that residues V215 and I216 in the linker are part of a hydrophobic pocket that includes W230 from helix 1 (Figure 6B). To test this premise, we carried out sedimentation velocity experiments to determine the oligomerization properties of mutants I216A and V215A/I216A. We previously showed that Rep68 sediments in two populations, a slow sedimenting population of monomer/dimers (~3 S) and a ~13 S population corresponding to oligomeric rings around (20). Figure 6D shows that when compared to WT, both mutant proteins have impaired oligomerization properties, with the double mutant forming only monomers. These results confirm that both linker residues are critical for the latch region to dock into the neighboring molecule. Looking back at the heptameric apo HD1 structure, we can also now distinguish density for the latch region that was not accounted for before due to the lower resolution of the map (Figure 4A, E). Consequently, the latch motif is flexible enough to stabilize ring structures with different stoichiometries producing rings that have a dynamic assembly/disassembly process. This property is illustrated by the variety of conformations that we see in the heptameric ring (Figure 6E) and the 3D variability analysis of the hexameric ring. Results from this analysis show that in one of the normal modes, a subunit coming in/out of the ring (Figure 6E, Supplementary Movie S2). Moreover, the 2D classes show a minor population of particles



**Figure 4.** Structure of apo Heptameric Helicase domain. (A) Local Cryo-EM density of top, side and bottom views. Cutoff of the side view of the HD showing the different diameters along the central channel (right panel). (B) Cryo-EM map and fitted Rep40 structure around the OD showing additional density that extends the N-terminal helix (in blue). (C) Ribbon representation of some of the potential interactions involved in the OD-OD interface. Positions of the side chains are based on the docked Rep40 X-ray model. The view is from the center of the channel. (D) Infectious particle assay of WT and the effect of N254A and S261A mutants. Data are from three independent experiments and are represented as the mean  $\pm$  SEM. (E) Side-bottom view of the HD1 Cryo-EM map displaying DNA density (yellow) and a docked poly-thymine (dT<sub>4</sub>) model. (F) Bottom view of the lower tier OD ring showing the position of residues R260 and K264 based on the docked Rep40 X-ray structure. (G) Cutoff showing surface representation of the electrostatics inside the central channel. Two electropositive rings (blue) are located in the AAA<sup>+</sup> domain at the PS1 $\beta$ H and at the OD. (H) Infectious particle assay of R260A and K264A mutants showing a drastic effect on the production of infectious particles compared to WT. Data are from three independent experiments and are represented as the mean  $\pm$  SEM. (I) Helicase assay. Representative gel of a set of duplicate experiments. Lane 1: WT, 0.25  $\mu$ g; lane 2: WT, 0.5  $\mu$ g; lane 3: WT, 1  $\mu$ g; lane 4: WT, no ATP; lane 5: fluorescein-ssDNA; lane 6: R260A no ATP; lane 7: R260A, 0.25  $\mu$ g; lane 8: R260A, 0.5  $\mu$ g; lane 9: R260A, 1  $\mu$ g; lane 10: K264A, 0.25  $\mu$ g; lane 11: K264A, 0.5  $\mu$ g; lane 12: K264A, 1  $\mu$ g; lane 13: K264A, no ATP.



**Figure 5.** Effect of ATP $\gamma$ S on Helicase Domain. (A) Two orthogonal views of the local HD1 cryo-EM map from the DOC and fitted heptameric model, blue density in the central channel is ssDNA. (B) Top (left) and side (right) views of the Rep68-ssAAVS1 cryo-EM density with fitted helicase model. Blue density in the central channel represents ssDNA. There is a clear induced curvature of the DNA molecule. (C) Representative Cryo-EM maps of three conformations of the Rep68-ssAAVS1 complex obtained from 3D variability analysis. The particles from the two right conformations were not used in the final model. (D) Superposition of APO (blue) and ATP $\gamma$ S (red) helicase domains showing the large conformational change of the AAA<sup>+</sup> domain. The HDs were superposed using only the ODs.



**Figure 6.** Cryo-EM structure of Rep68-ssAAVS1-ATP $\gamma$ S hexameric complex. **(A)** Top and side views of the cryo-EM hexameric complex with fitted helicase model. **(B)** Details of the flexible latch (pink) docking into a pocket of the neighboring subunit. **(C)** Ribbon representation of the ODs showing the latch of each subunit in the hexameric ring. **(D)** Sedimentation velocity analysis of Rep68 wt (black line) and the I216A (blue line) and V215A/I216A (red line) mutants. All proteins were analyzed at a concentration of 0.25 mg/ml, as described in Material and Methods. **(E)** Cryo-EM maps of three normal models obtained through 3D variability analysis. There is an in/out movement of one of the subunits. **(F)** Representative 2D classes of the hexamer and its dynamic behavior. The classes with loose and missing sixth subunit represent 6.1% and 3.1% of all hexameric particles and were not included in the final reconstruction.

with only five subunits (Figure 6F). Taken together, our results imply that upon ATP binding and/or hydrolysis, there is formation of hexameric rings, possibly by disassembly of heptamers.

## DISCUSSION

The results presented in this report illustrate the structural basis conferring AAV Rep proteins the conformational plasticity necessary to acquire multiple oligomeric states upon binding to different types of DNA substrates. This oligomeric malleability originates from three key features: the presence of multiple oligomerization interfaces located in three different domains, the existence of several protein–DNA interaction motifs and the fact that each domain can form different stoichiometric oligomers independent of each other. Thus, the OBDs associate as octamers and the HDs tend to form heptamers and hexamers. The active participation of the OBD in oligomerization was unexpected, as no reports from other members of the SF3 and HUH families having this property have been described previously. As shown by our AUC and X-ray studies, the propensity to form octameric rings is inherent to the OBD, as it can form octamers at high concentrations ( $>200\ \mu\text{M}$ ), illustrating that this oligomerization interface is not particularly robust. However, in the context of the LReps and under the right conditions, the formation of octamers occurs naturally (20). It is clear from the cryo-EM and X-ray structures that the formation of the double-octameric complex occurs through ssDNA bridging of two octameric complexes requiring a ssDNA molecule with a minimum size of 17 nucleotides in length and containing a stretch of at least three pyrimidines, preferably thymines. At the present time, we don't know if the DOC participates in any of the processes during the life cycle of the AAV. However, a simple inspection of the AAV genome sequence shows that sections with three and four consecutive thymines occur at least 50 and 12 times, respectively, and segments with successive thymines and cytosines at higher frequencies. This suggests that, in principle, there are enough sites in the AAV genome for the formation of the double-octameric complex. Still, it is also possible that the DOC-structure only forms *in vitro* as a result of the conditions used in our studies, particularly the use of a dT<sub>25</sub> ssDNA. Nevertheless, the use of poly-dT has permitted us to trap the Rep68 octameric complex in a conformation that may be difficult to capture due to the inherently dynamic nature of the system. This dynamic character can be seen in the cryo-EM structures of the heptameric Rep68 with ssAAVS1 (Figure 5) and Rep68–dT<sub>25</sub> hexamer (Figure 6), where the OBD domains by being not engaged in octameric interface formation, are not resolvable due to their high mobility.

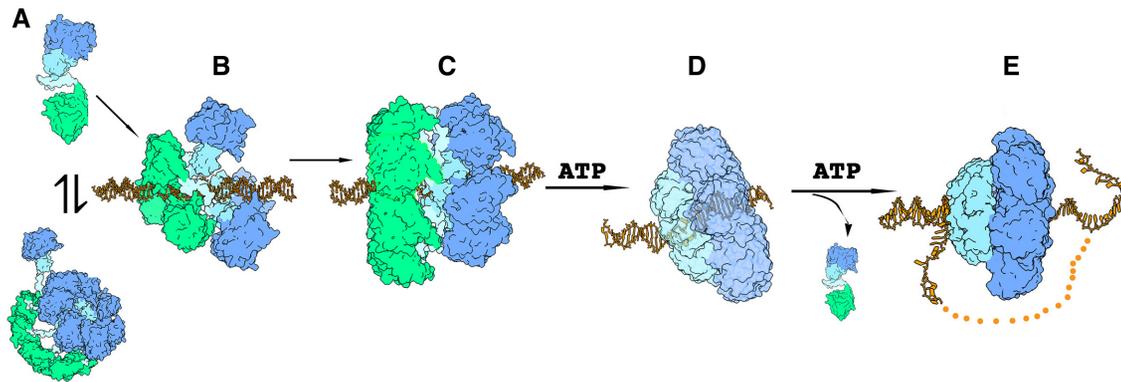
The question remains as to what the functional significance of the OBD octameric structure is. One of the most intriguing features of the DOC is that the interface utilizes the same motifs—and some of the same residues—as those that participate in the specific recognition of RBS dsDNA sequences found at the AAV origin of replication, p5 promoter, and AAVS1 integration site. Therefore, it is tempting to postulate that after the initial binding of Rep to RBS sites, the interface switches from a DNA recognition mode

into an oligomerization mode. This arrangement puts the OBD ring in a conformation such that the HUH nuclease catalytic core is located in the exterior face of the ring, thus preventing any unintended hydrolysis of the DNA backbone. Further studies will be needed to explore this hypothesis.

We had previously shown that the interdomain linker plays an active part in Rep68 oligomerization; the work presented here provides the structural determinants of such function (22). The cryo-EM structures illustrate that as predicted, the linker is partly folded and extends the N-terminal helix of the HD by  $\sim 6$  residues (18,22). More importantly, five additional linker residues preceding this N-terminal helix form a flexible latch that docks into the OD of a neighboring subunit (Figure 7C). Remarkably, the same interaction can be observed in both the heptameric and the hexameric complexes. We propose that this region (aa 215–220) acts as a flexible latch to induce oligomerization and stabilize the ring structures. Is this motif unique to Rep proteins? Analysis of the X-ray structures and biochemical studies of SV40-Tag show that it assembles as hexamers without the need of a linker latch region (38,39). Similarly, the first E1 helicase X-ray structure lacked the equivalent linker residues but still was able to form hexameric rings (40). However, subsequent structures with E1 constructs that included the linker region show a similar flexible latch that interacts with the neighboring subunit, although in opposite direction (41). Thus, although the latch motif may not promote oligomerization in other SF3 helicases, it is absolutely necessary in AAV Rep proteins and can be considered an integral part of the helicase domain.

Another unexpected feature of the apo Rep68 structure, when compared to other SF3 family members, is that the AAA<sup>+</sup> domain does not participate in the oligomerization process. Thus, the domain that dictates HD oligomerization resides solely in the OD, which produces a heptameric ring. Additionally, another significant difference between Rep and other SF3 helicases is the extent of conformational change that is triggered upon nucleotide binding. In Rep68, the AAA<sup>+</sup> domain undergoes a large rigid body rotation with respect to the apo structure of  $\sim 44^\circ$  while in E1 and SV40-Tag, this rotation is of only  $4.8^\circ$  and  $12.8^\circ$  respectively (Supplementary Figure S11). The magnitude of the conformational change in Rep proteins is in part due to the configuration of the region that links the OD to the AAA<sup>+</sup> domain. In E1 and SV40-tag, this region consists of a long  $\alpha$ -helix that folds into the AAA<sup>+</sup> domain. In contrast, Rep proteins have a long linker with a small  $\alpha$ -helix of only four residues that interact loosely with the AAA<sup>+</sup> domain (Supplementary Figure S11). Unfortunately, the current resolution of our cryo-EM data prevents further analysis of the interactions leading to this large conformational change and will require further structural studies.

From a functional perspective, the accepted model of DNA translocation by SF3 helicases during DNA unwinding does not require a large conformational change of the entire AAA<sup>+</sup> domain. Instead, alternating conformations of the PS1 $\beta$ H during ATP binding and hydrolysis are sufficient to drive the process. Therefore, the question remains as to the functional significance of the large conformational change in Rep proteins. A simple explanation is that both



**Figure 7.** Model for assembly and DNA melting on dsDNA by AAV Rep proteins. (A) Rep68 is in equilibrium between monomers and higher order oligomers including octamers. (B) Initial assembly of the complex is directed by the OBDs that specifically recognizes GCTC repeats found in RBS dsDNA sites imparting directionality to the complex. (C) A HD heptameric ring assembles around DNA and (D) upon binding ATP generates a deformation that produces strand separation, the role of the ODs in this step is not known, if any. (E) Generation of ssDNA leads to a conversion from heptamer to hexamer with the expulsion of a Rep monomer. The OBD domains are not shown in steps (D) to (E) for simplification.

E1 and SV40 are mainly processive helicases that need to translocate on DNA efficiently while AAV LReps are used in multiple processes involved in DNA structural reconfiguration and melting. In this context, the large conformational change seen in Rep68 may induce DNA distortions leading eventually to its melting. The degree of conformational change is similar in both heptameric and hexameric complexes and also involves the generation of a new DNA contact point at the PS1 $\beta$ H region. As a result, Rep68 ring structures contact DNA inside the channel at two locations separated by  $\sim 4$ – $5$  bp, a property that again is not shared with other SF3 helicases, which contact DNA using only residues in the PS1 $\beta$ H region (40). The large conformational change may also explain how a mostly monomeric Rep40 may be able to unwind DNA without being a processive ring helicase, as the conformational change may be sufficient to unwind DNA without the need of forming a ring. Additional research will be needed to explore this hypothesis.

While the majority of AAA<sup>+</sup> proteins, including most SF3 helicases, form stable oligomers irrespective of nucleotide binding, a small fraction of their members display a dynamic oligomeric behavior in a concentration and nucleotide-dependent manner (42). Our results show that Rep68 HD is of the latter type as its oligomerization is highly dependent on concentration, presence of nucleotide and DNA substrate. Moreover, the Rep HD is unique among SF3 family members by forming heptameric and hexameric rings. However, this property is not exclusive to AAV LReps, as some SF4 helicases have been reported as having this capability. For instance, gene 4 from bacteriophage T7 and the mitochondrial replicative helicase Twinkle also form mixtures of hexamers/heptamers (43–46). Compared to other helicases, what advantage does this property bring to AAV Rep proteins? Mechanistically, this ability might be a reflection of the multifunctionality of Rep proteins that are required to participate in multiple DNA transactions during the AAV life cycle, including DNA replication, transcription regulation, and genome packaging. Nevertheless, a common step in all these pro-

cesses remains the assembly of a Rep complex after recruitment to the RBS site. In this context, the transition from heptameric to hexameric rings may be significant during the loading of Rep proteins onto dsDNA. Generally, the loading of a replicative ring helicase onto DNA requires the help of a helicase loader to open up the ring, thus allowing access of DNA. However, it has been hypothesized that in some helicase systems, an already open heptameric ring can easily assemble to either single or double-stranded DNA without the need of an additional factor (47–51). Once loaded, it switches to a processive hexameric form through the loss of one of the subunits in a process that is NTP and/or DNA dependent (52). A similar process seems fitting for Rep mediated processes. We have previously shown that Rep68 forms heptameric rings when bound to an AAVS1 dsDNA site (17). In addition, during AAV DNA replication, a necessary step during the terminal resolution reaction is the recognition of an RBS site, the formation of a Rep-DNA complex, and the eventual melting of the dsDNA. Based on our studies, we propose a mechanism where, LReps initiate the assembly process by binding to RBS sites where the OBDs recognizes GCTC sites, in this way conferring directionality to the complex (Figure 7). This initial step may involve 2–3 LReps molecules (Figure 7B) that will eventually recruit additional LReps to further assemble into a heptameric complex (Figure 7C). Some of the OBDs may remain bound to the RBS site and may contribute to subsequent steps. As a heptamer, the diameter of the central channel is wide enough to accommodate a dsDNA molecule, while in the hexameric form, this is not possible due to compression of the diameter of the channel to  $\sim 13$  Å (Figures 5 and 6). Could DNA melting occur during the transition to a hexameric ring? Our data show that DNA deformation occurs in the heptameric complex, which could induce further structural deformations leading to DNA melting (Figure 7D, E). We hypothesize that the asymmetry of the heptameric complex is one of the drivers of DNA deformation as the unbalanced neutralization of charges of the phosphate backbone is a well-known cause of DNA bending (53). Furthermore, the presence of two

contact points in the DNA and the large conformational change upon nucleotide binding can produce additional deformation and DNA melting. While future studies will be needed to verify this model, our data has revealed important insights into the mode of action of the AAV Rep proteins.

## DATA AVAILABILITY

All cryo-EM maps and main chain atomic coordinates have been deposited to EMD. Densities for the DOC apo and ATP $\gamma$ S-bound structures are EMD-22451, EMD-22452 and EMD-22453 respectively. Rep68-ssAAVS1-ATP $\gamma$ S density is EMD-22454 and PDB: 7JSH. Rep68-ssdT<sub>25</sub> hexamer density is EMD-22455 and PDB:7JSI. X-ray structure coordinates for AAV2 OBD-ssRBS complex is PDB:6X88.

## SUPPLEMENTARY DATA

[Supplementary Data](#) are available at NAR Online.

## ACKNOWLEDGEMENTS

We would like to thank Ed Eng, Venkata Dandey, Elina Kopylov and the outstanding staff from the National Center for Cryo-EM Access and Training (NCCAT). We would like to acknowledge Lauren Hales Beck, Nancy Meyer from the Pacific Northwest Cryo-EM Center (PNCC) supported by NIH grant U24GM129547.

*Author Contributions:* V.S., E.H., C.R.E., designed research; V.S., R.J., F.M., B.V., FZP, KS, C.D., M.E., K.D. and C.R.E. performed research; V.S. and C.R.E. wrote the paper with all authors providing comments or revisions.

## FUNDING

C.R.E. is supported by National Institute of Health (NIH) [R01 GM124204]; E.H. is supported by a start-up grant from KU Leuven and research grant from the Flemish Research Council [FWO-G0C3220N]; a portion of this research was carried out at the NCCAT and the Simons Electron Microscopy Center located at the New York Structural Biology Center, supported by the NIH Common Fund Transformative High Resolution Cryo-Electron Microscopy program [U24 GM1 129539]; Simons Foundation [SF349247]; NY State; portion of this research was supported by NIH grant [U24GM129547] and performed at the PNCC at OHSU and accessed through EMSL [grid.436923.9]; DOE Office of Science User Facility sponsored by the Office of Biological and Environmental Research. Funding for open access charge: National Institute of Health.

*Conflict of interest statement.* Els Henckaerts has a sponsored research agreement with Handl Therapeutics.

## REFERENCES

- Chiorini, J.A., Wiener, S.M., Owens, R.A., Kyostio, S.R.M., Kotin, R.M. and Safer, B. (1994) Sequence requirements for stable binding and function of Rep68 on the adenovirus type 2 inverted terminal repeats. *J. Virol.*, **68**, 7448–7457.
- McCarthy, D.M., Ryan, J.H., Zolotukhin, X.Z. and Muzyczka, N. (1994) Interaction of the adeno-associated virus Rep protein with a sequence within the A palindrome of the viral terminal repeat. *J. Virology*, **68**, 4998–5006.
- Prasad, K.M. and Trempe, J.P. (1995) The adeno-associated virus Rep78 protein is covalently linked to viral DNA in a preformed virion. *Virology*, **214**, 360–370.
- Smith, R.H. and Kotin, R.M. (1998) The Rep52 gene product of adeno-associated virus is a DNA helicase with 3'-to-5' polarity. *J. Virol.*, **72**, 4874–4881.
- Dubielzig, R., King, J.A., Weger, S., Kern, A. and Kleinschmidt, J.A. (1999) Adeno-associated virus type 2 protein interactions: formation of pre-encapsidation complexes. *J. Virol.*, **73**, 8989–8998.
- Zhou, X., Zolotukhin, I., Im, D.S. and Muzyczka, N. (1999) Biochemical characterization of adeno-associated virus rep68 DNA helicase and ATPase activities. *J. Virol.*, **73**, 1580–1590.
- Beaton, A., Palumbo, P. and Berns, K.I. (1989) Expression from the adeno-associated virus p5 and p19 promoters is negatively regulated in trans by the rep protein. *J. Virol.*, **63**, 4450–4454.
- King, J.A., Dubielzig, R., Grimm, D. and Kleinschmidt, J. A. (2001) DNA helicase-mediated packaging of adeno-associated virus type 2 genomes into preformed capsids. *EMBO J.*, **20**, 3282–3291.
- Im, D.S. and Muzyczka, N. (1990) The AAV origin binding protein Rep68 is an ATP-dependent site-specific endonuclease with DNA helicase activity. *Cell*, **61**, 447–457.
- Yang, Q., Kadam, A. and Trempe, J.P. (1992) Mutational analysis of the adeno-associated virus rep gene. *J. Virol.*, **66**, 6058–6069.
- Owens, R.A., Weitzman, M.D., Kyostio, S.R. and Carter, B.J. (1993) Identification of a DNA-binding domain in the amino terminus of adeno-associated virus Rep proteins. *J. Virol.*, **67**, 997–1005.
- James, J.A., Escalante, C.R., Yoon-Robarts, M., Edwards, T.A., Linden, R.M. and Aggarwal, A.K. (2003) Crystal structure of the SF3 helicase from adeno-associated virus type 2. *Structure*, **11**, 1025–1035.
- James, J.A., Aggarwal, A.K., Linden, R.M. and Escalante, C.R. (2004) Structure of adeno-associated virus type 2 Rep40-ADP complex: insight into nucleotide recognition and catalysis by superfamily 3 helicases. *Proc. Natl. Acad. Sci. U.S.A.*, **101**, 12455–12460.
- James, A.B., Ronning, D.R., Kotin, R.M. and Dyda, F. (2002) Structural unity among viral origin binding proteins: crystal structure of the nuclease domain of adeno-associated virus Rep. *Mol. Cell*, **10**, 327–337.
- Hickman, A.B., Ronning, D.R., Perez, Z.N., Kotin, R.M. and Dyda, F. (2004) The nuclease domain of adeno-associated virus rep coordinates replication initiation using two distinct DNA recognition interfaces. *Mol. Cell*, **13**, 403–414.
- Hickman, A.B. and Dyda, F. (2005) Binding and unwinding: SF3 viral helicases. *Curr. Opin. Struct. Biol.*, **15**, 77–85.
- Musayev, F.N., Zarate-Perez, F., Bishop, C., Burgner, J.W. 2nd and Escalante, C.R. (2015) Structural insights into the assembly of the Adeno-associated virus type 2 Rep68 protein on the integration site AAVS1. *J. Biol. Chem.*, **290**, 27487–27499.
- Musayev, F.N., Zarate-Perez, F., Bardelli, M., Bishop, C., Saniev, E.F., Linden, R.M., Henckaerts, E. and Escalante, C.R. (2015) Structural studies of AAV2 Rep68 reveal a partially structured linker and compact domain conformation. *Biochemistry*, **54**, 5907–5919.
- Yoon-Robarts, M., Blouin, A.G., Bleker, S., Kleinschmidt, J.A., Aggarwal, A.K., Escalante, C.R. and Linden, R.M. (2004) Residues within the B' motif are critical for DNA binding by the superfamily 3 helicase Rep40 of adeno-associated virus type 2. *J. Biol. Chem.*, **279**, 50472–50481.
- Zarate-Perez, F., Mansilla-Soto, J., Bardelli, M., Burgner, J.W. 2nd, Villamil-Jarauta, M., Kekilli, D., Samso, M., Linden, R.M. and Escalante, C.R. (2013) Oligomeric properties of adeno-associated virus Rep68 reflect its multifunctionality. *J. Virol.*, **87**, 1232–1241.
- Ogura, T. and Wilkinson, A.J. (2001) AAA+ superfamily ATPases: common structure-diverse function. *Genes Cells*, **6**, 575–597.
- Zarate-Perez, F., Bardelli, M., Burgner, J.W. 2nd, Villamil-Jarauta, M., Das, K., Kekilli, D., Mansilla-Soto, J., Linden, R.M. and Escalante, C.R. (2012) The interdomain linker of AAV-2 Rep68 is an integral part of its oligomerization domain: role of a conserved SF3 helicase residue in oligomerization. *PLoS Pathog.*, **8**, e1002764.
- Mansilla-Soto, J., Yoon-Robarts, M., Rice, W.J., Arya, S., Escalante, C.R. and Linden, R.M. (2009) DNA structure modulates

- the oligomerization properties of the AAV initiator protein Rep68. *PLoS Pathog.*, **5**, e1000513.
24. Bardelli, M., Zarate-Perez, F., Agundez, L., Jolinon, N., Linden, R.M., Escalante, C.R. and Henckaerts, E. (2017) Analysis of replicative intermediates of Adeno-associated virus through hirt extraction and Southern Blotting. *Bio. Protoc.*, **7**, e2271.
  25. Schuck, P. (2000) Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophys. J.*, **78**, 1606–1619.
  26. Zheng, S.Q., Palovcak, E., Armache, J.P., Verba, K.A., Cheng, Y. and Agard, D.A. (2017) MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods*, **14**, 331–332.
  27. Rohou, A. and Grigorieff, N. (2015) CTFFIND4: fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.*, **192**, 216–221.
  28. Scheres, S.H. (2012) RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.*, **180**, 519–530.
  29. Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C. and Ferrin, T.E. (2004) UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.*, **25**, 1605–1612.
  30. Emsley, P. and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. D. Biol. Crystallogr.*, **60**, 2126–2132.
  31. Adams, P.D., Grosse-Kunstleve, R.W., Hung, L.W., Ioerger, T.R., McCoy, A.J., Moriarty, N.W., Read, R.J., Sacchettini, J.C., Sauter, N.K. and Terwilliger, T.C. (2002) PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr. D. Biol. Crystallogr.*, **58**, 1948–1954.
  32. Kleywegt, G.J.A.J., T.A. (1994) A super position. *CCP4/ESF-EACBM Newsllett. Protein Crystallogr.* **31**, 9–14.
  33. Winn, M.D., Ballard, C.C., Cowtan, K.D., Dodson, E.J., Emsley, P., Evans, P.R., Keegan, R.M., Krissinel, E.B., Leslie, A.G.W., McCoy, A. et al. (2011) CCP4: overview of the CCP4 suite and current developments. *Acta Crystallogr. D. Biol. Crystallogr.*, **67**, 235–242.
  34. Brookes, E., Demeler, B., Rosano, C. and Rocco, M. (2010) The implementation of SOMO (SOLUTION MOdeller) in the UltraScan analytical ultracentrifugation data analysis suite: enhanced capabilities allow the reliable hydrodynamic modeling of virtually any kind of biomacromolecule. *Eur. Biophys. J.*, **39**, 423–435.
  35. Brookes, E. and Rocco, M. (2018) Recent advances in the UltraScan SOLUTION MOdeller (US-SOMO) hydrodynamic and small-angle scattering data analysis and simulation suite. *Eur. Biophys. J.*, **47**, 855–864.
  36. Punjani, A., Rubinstein, J.L., Fleet, D.J. and Brubaker, M.A. (2017) cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods*, **14**, 290–296.
  37. Punjani, A. and Fleet, D.J. (2020) 3D variability analysis: directly resolving continuous flexibility and discrete heterogeneity from single particle cryo-EM images. bioRxiv doi: <https://doi.org/10.1101/2020.04.08.032466>, 09 April 2020, preprint: not peer reviewed.
  38. Li, D., Zhao, R., Lilyestrom, W., Gai, D., Zhang, R., DeCaprio, J.A., Fanning, E., Jochimiak, A., Szakony, G. and Chen, X.S. (2003) Structure of the replicative helicase of the oncoprotein SV40 large tumor antigen. *Nature*, **423**, 512–518.
  39. Wessel, R., Schweizer, J. and Stahl, H. (1992) Simian virus 40 T-antigen DNA helicase is a hexamer which forms a binary complex during bidirectional unwinding from the viral origin of DNA replication. *J. Virol.*, **66**, 804–815.
  40. Enemark, E.J. and Joshua-Tor, L. (2006) Mechanism of DNA translocation in a replicative hexameric helicase. *Nature*, **442**, 270–275.
  41. Sanders, C.M., Kovalevskiy, O.V., Sizov, D., Lebedev, A.A., Isupov, M.N. and Antson, A.A. (2007) Papillomavirus E1 helicase assembly maintains an asymmetric state in the absence of DNA and nucleotide cofactors. *Nucleic Acids Res.*, **35**, 6451–6457.
  42. Sysoeva, T.A. (2017) Assessing heterogeneity in oligomeric AAA+ machines. *Cell. Mol. Life Sci.*, **74**, 1001–1018.
  43. Singleton, M.R., Sawaya, M.R., Ellenberger, T. and Wigley, D.B. (2000) Crystal structure of T7 gene 4 ring helicase indicates a mechanism for sequential hydrolysis of nucleotides. *Cell*, **101**, 589–600.
  44. Toth, E.A., Li, Y., Sawaya, M.R., Cheng, Y. and Ellenberger, T. (2003) The crystal structure of the bifunctional primase-helicase of bacteriophage T7. *Mol. Cell*, **12**, 1113–1123.
  45. Crampton, D.J., Mukherjee, S. and Richardson, C.C. (2006) DNA-induced switch from independent to sequential dTTP hydrolysis in the bacteriophage T7 DNA helicase. *Mol. Cell*, **21**, 165–174.
  46. Fernandez-Millan, P., Lazaro, M., Cansiz-Arda, S., Gerhold, J.M., Rajala, N., Schmitz, C.A., Silva-Espina, C., Gil, D., Bernado, P., Valle, M et al. (2015) The hexameric structure of the human mitochondrial replicative helicase Twinkle. *Nucleic Acids Res.*, **43**, 4284–4295.
  47. Funnell, B.E., Baker, T.A. and Kornberg, A. (1987) In vitro assembly of a prepriming complex at the origin of the Escherichia coli chromosome. *J. Biol. Chem.*, **262**, 10327–10334.
  48. Kobori, J.A. and Kornberg, A. (1982) The Escherichia coli dnaC gene product. III. Properties of the dnaB-dnaC protein complex. *J. Biol. Chem.*, **257**, 13770–13775.
  49. Kobori, J.A. and Kornberg, A. (1982) The Escherichia coli dnaC gene product. II. Purification, physical properties, and role in replication. *J. Biol. Chem.*, **257**, 13763–13769.
  50. Kobori, J.A. and Kornberg, A. (1982) The Escherichia coli dnaC gene product. I. Overlapping of the dnaC proteins of Escherichia coli and Salmonella typhimurium by cloning into a high copy number plasmid. *J. Biol. Chem.*, **257**, 13757–13762.
  51. Soultanas, P. (2012) Loading mechanisms of ring helicases at replication origins. *Mol. Microbiol.*, **84**, 6–16.
  52. Crampton, D.J., Ohi, M., Qimron, U., Walz, T. and Richardson, C.C. (2006) Oligomeric states of bacteriophage T7 gene 4 primase/helicase. *J. Mol. Biol.*, **360**, 667–677.
  53. Strauss, J.K. and Maher, L.J. 3rd (1994) DNA bending by asymmetric phosphate neutralization. *Science*, **266**, 1829–1834.